

# Coherence and Cohesion in EFL Learners' Text -An Analysis of One EFL Learner's Writing-

Aiko Ono

## 1. Introduction

In Japan's English education for junior high and high school, teachers teach practical English focusing on all four skills: reading, listening, writing and speaking. Benesse Educational Research and Development Institute (2009) conducted the first survey of junior high school English. In the survey, 72 % of second-year students answered "yes" to a question that "Is it difficult to write sentences in English?". This was the highest value among all skills. One of the four skills, writing, is particularly complicated and difficult for learners to acquire. It might be hard for them to write their own opinions logically and they might have trouble making sentences with coherence and cohesion such as connecting ideas and using connectives. According to the Ministry of Education, Culture, Sports, Science and Technology (MEXT hereafter, 2008), the course of study for junior high school foreign languages says junior high school students are required to *write correct sentences paying attention to connection between words* in writing activity. It indicates that EFL learners might think it is difficult to write sentences with coherence. Under these circumstances, the way of teaching coherence and cohesion needs to be reconsidered. This paper focuses on coherence and cohesion in one EFL learner's text. In this paper, coherence is defined as a bridge between sentences (Bain, 1890 as cited in Lee, 2002), and cohesion is defined as a relation of meaning that exists in text (Halliday & Hasan, 1976). Coherence can be said to be internal to the text, adding to our understanding of what a text is (Lee, 2002). On the other hand, cohesion is surface relation in a text and it contributes to develop understanding of meaning in text (Kudo, 2009). One example of cohesion is

reiteration (Halliday & Hasan, 1976). By repeating key ideas using synonyms, hypernyms, and hyponyms the writer can create lexical cohesion which will contribute to overall coherence. Equipping teachers with such knowledge of coherence and cohesion would help them teach writing to EFL learners. This paper identifies problems with coherence and cohesion in learners' writing and provides a solution in the form of instruction.

One problem is that a learner who knows how to structure a paragraph may not know how to *connect* sentences. In other words, learners might not know how to connect ideas at discourse level even though they learn how to write long paragraphs. EFL learners tend to start sentences with "I" in English writing (e.g., Oi, 2008 and Kudo, 2011). They write their paragraphs without caring about topical development. In his study of cohesion in Japanese EFL students' writing, Kudo (2011) examines the study conducted by National Institute for Educational Policy Research (2005). In the study by National Institute for Educational Policy Research (2005), junior high school students repeated sentences like "I like dogs", "I like banana" and "I like cat" when they were asked to write over four sentences about given topics. The National Institute for Educational Policy Research established a cohesive standard which was students wrote over four sentences and the sentences were well connected. When it assessed using the cohesive standard, the students who met the standard were less than 40%. Additionally, a survey by National Institute for Educational Policy Research (2007), only 25% of high school students met a cohesive standard. The cohesive standard was that students wrote over four sentences and the structure of writing was regarded as appropriate one. In English classes nowadays, almost all teachers tell learners how to write paragraphs in a hamburger-style. Oi (2008) presents introduction of teaching plans with hamburger-style. The hamburger-style is about putting the topic sentence on the top, supporting sentences in the middle and the concluding sentence on the bottom. Yet, learners might not write with coherence or cohesion even when they know appropriate organizations of essays, because they may not know how to connect ideas.

In the course of study for foreign language for high school, it says learners are required to write with coherence and also use conjunctions, which is an aspect of cohesion (MEXT, 2009). However, it does not provide sufficient examples of

coherence and cohesion. Without concrete examples, as will be shown later many teachers might not know how to teach coherence and cohesion. That is because, some teachers might not teach learners how to *connect* ideas. It is important for them not only to know coherence and cohesion but also to use and apply knowledge about coherence and cohesion in writing class. As an EFL learner, the author did not learn how to connect ideas in writing when studying at junior high and high school. When teachers refer to coherence, they frequently use terminology unfamiliar to learners such as “the ideas don’ t hang together” or “your writing is not coherent” (Lee, 2002 p.136). Such advice is vague, so learners do not seem to understand specific problems the teacher sees. That is, the teacher and learners are not communicating the idea of coherence and cohesion very well. To improve learners’ writing skills, the instructional approach needs to be changed.

Based on these problems, what kind of instructional tool, if any, could aid teachers provide better feedback on learners’ writing? This paper analyzed the author’s writing and created a checklist (see Appendix 1) and a list of common errors (see Appendix 2) to help teachers and learners detect problems in writing related to coherence and cohesion. In order to demonstrate what problems an EFL learner might have, this paper first analyzed the author’s past essays. In the analysis, Coh-Metrix and Theme-Rheme structure were used to see changes between two essays. Coh-Metrix was used to see cohesion in text by quantifying characteristic of text such as word information and readability. Theme-Rheme structure was used to see degrees of Thematic progression in texts. This paper adopts Halliday’s (1985) definition of the Theme which is given information or “the point of departure” of a message. Rheme is remainder of a message in a clause (Halliday, 1985, p.38). The difference between them is simple: Theme contains old or familiar information and Rheme contains new or unfamiliar information. Thematic progression is defined as “the exchange of information between successive Theme and Rheme pairings in a text” (Wang, 2007, p.168). It has various definitions depending on types of writing. Thematic progression in academic text needs to have cross-referential links from the Rheme of one clause to the Theme of the next clause (Fries, 1983 as cited in Wang, 2007). In other words, the first-appearing Rheme becomes the Theme in the next clause. This Thematic progression gives readers

orientation (Wang, 2007, p.169).

The checklist (Appendix 1) and the list of common errors (Appendix 2) were created to offer a solution to improve teachers' feedback regarding coherence and cohesion. The checklist was created based on Halliday & Hasan (1976), Halliday (1985) and Eggins (1994), and it arranged coherence and cohesion (especially Theme-Rheme structure) for classroom use. The list of common errors was created based on Bloor & Bloor categories (1992 as cited in Wang, 2007), Eggins (1994) and the author's writing. There are three problems identified by Bloor & Bloor (1992) such as brand new Theme, double Rhemes and empty Rheme. As described above, a Theme usually has old or familiar information and a Rheme has new or unfamiliar information. The problem of brand new Theme is putting new information in the Theme position. Inexperienced writers frequently commit this error. The problem of double Rhemes is having two Rhemes in a clause, one of which has no previous mention. The problem of empty Rheme is having no new information in the Rheme position. As mentioned, a Rheme has to present new or unfamiliar information. Each example is shown Appendix 2. Both of the materials, the checklist and the list of common errors were written English and Japanese because they can be used by learners and teachers, both Japanese and non-Japanese.

## **2. The Study**

This study employs Coh-Metrix and Theme-Rheme structure for the analysis of the author's past essays to find changes or improvements, if any, in coherence and cohesion.

### **2.1 The author's profile**

The author is a 4th year university student who went through the public school system in Japan. The author has grade 2 of the EIKEN Test in Practical English Proficiency and a TOEIC score of 725 as of 2017. In addition, the author took private English lessons to study all skills (reading, listening, writing, and speaking) for 7

years since the 6th grade. Writing is one of the required subjects for students who major in English at a private four-year university, so the author studied it every semester. However, the author personally feels that the authentic way of writing was finally taught in her 2nd year at the private four-year university when a dedicated American instructor taught writing focusing on connecting ideas and sentences.

## **2.2 Materials**

The essays analyzed were written in February 2015 and July 2016. The former was written using a Japanese-English dictionary when the author was a first-year student on a four-month study abroad program in the U.S. with a TOEIC score of 650. The content of the writing was improvements and complaints about the four-month study abroad program in the U.S. Before this 2015 essay was written, the author learned how to write essay such as structure of essay. The latter was written using a Japanese-English dictionary when the author was in third year at the university with a TOEIC score of 670. The content was a review of critical on a research paper about a change in speaking performance of Japanese high school students. As an EFL writer, the author learned how to connect ideas in writing class before the 2016 essay was written. Both of those essays were written with no time limit. The author's TOEIC score was used to see a relation between English proficiency and quality of writing.

## **2.3 Instruments**

This paper employed Coh-Metrix and Theme-Rheme structure (Wang, 2007). This study used Coh-Metrix 3.X version to analyze the author's essays for readability, word information and cohesion. The Coh-Metrix is a tool that can quantify characteristic of words in texts. This tool can tell computational cohesion and analyze lexical diversity, connective incidence, word information and readability by numerical value. To see surface changes in text in terms of words, phrases and connectives incidence, this paper employed Coh-Metrix. The Theme-Rheme structure demonstrates whether the text has Thematic progression or not. The Thematic progression in academic text needs to have high incidence of cross-referential links between Theme and Rheme (Wang, 2007, p.168). That is to say, the first-appearing Rheme needs to become Theme in the

next clause. As described above, a Theme usually has old or familiar information and a Rheme has new or unfamiliar information.

## 2.4 Procedures

### 2.4.1 Coh-Metrix

First, both of the essays were analyzed using the Coh-Metrix web tool. Second the results about basic information (e.g., the number of words and sentences), and cohesion were selected by the author for analysis. Then the results for other elements (e.g., narrativity, negation and LSA overlap) were excluded because of irrelevance to this study. Lastly, results were shown on Table 1.

### 2.4.2 Theme-Rheme structure

Firstly, five sentences were excerpted from a paragraph of the 2015 essay and the 2016 essay. Second, each clause was divided into a Theme and a Rheme. Then, the Thematic progression was checked. In this study, only the five sentences were analyzed because the whole essay was lengthy.

## 2.5 Result and Discussion of the Coh-Metrix analysis

By using Coh-Metrix, the 2015 essay and the 2016 essay were analyzed in detail. Table 1 shows the results of both of essays. Detailed descriptions of each label can be found in Appendix 3.

Table 1 Results of both of the 2015 essay and the 2016 essay

	Label	2015	2016	Description
1	DESPC	5.0	11.0	Paragraph count, number of paragraphs
2	DESSC	82.0	70.0	Sentence count, number of sentence
3	DESWC	752.0	986.0	Word count, number of words
4	DESPL	7.5	6.4	Paragraph length, number of sentences in a paragraph, mean
5	LDTTRa	0.4	0.3	Lexical diversity, type-token ratio, all words
6	CNCCausa	42.6	27.4	Causal connectives incidence
7	CNCLogic	49.2	46.7	Logical connectives incidence

8	CNCADC	12.0	20.3	Adversative and contrastive connectives incidence
9	CNCTemp	14.6	29.4	Temporal connectives incidence
10	CNCAdd	49.2	73.0	Additive connectives incidence
11	CNCPos	0.0	0.0	Positive connectives incidence
12	CNCNeg	0.0	0.0	Negative connectives incidence
13	SMTEMP	0.8	0.9	Temporal cohesion, tense and aspect repetition, mean
14	WRDNOUN	300.5	318.5	Noun incidence
15	WRDVERB	138.3	119.7	Verb incidence
16	WRDADJ	85.1	101.4	Adjective incidence
17	WRDADV	57.2	42.6	Adverb incidence
18	WRDPRO	81.1	25.4	Pronoun incidence
19	WRDAOAc	362.6	405.4	Age of acquisition for content words, mean
20	WRDFAMc	587.3	576.9	Familiarity for content words, mean
21	RDFRE	66.9	47.4	Flesch Reading Ease
22	RDFKGL	6.2	10.2	Flesch-Kincaid Grade level

This analysis was conducted using Coh-Metrix to see the essays' cohesion regarding the number of connectives and word-incidences in the analysis of the author's past writing. According to Table 1, there were similarities in some items including logical connectives, temporal cohesion and noun and adverb incidences, but differences were also found in the grade level and readability. The 2015 essay consisted of 5 paragraphs, 82 sentences and 752 words. The 2016 essay consisted of 11 paragraphs, 70 sentences and 986 words. An average number of words in a sentence increased; that of the 2015 essay was 9 words and that of the 2016 essay was 14 words. It indicates the 2016 essay had longer sentences with more connectives compared to the 2015 essay. Incidences of connectives were approximately twice as high as those of the 2015 essay, including adversative (e.g., yet), contrastive (e.g., on the other hand), temporal (e.g., before) and additive (e.g., plus). Age of acquisition for content words (WRDAOAc, no.19) showed that the 2015 essay's numerical value was 362. As shown in Appendix 3, the numerical value of "milk", "pony" and "smile" was 202. The numerical value of both Flesch reading ease (RDFRE, no.21) were 66 in the 2015 essay and 47 in the 2016

essay. Higher scores show easier reading, so the 2015 essay was easier to read than the 2016 essay. The grade level (RDFKGL, no. 22) of the 2015 essay was 6th (senior at elementary schools) and the 2016 essay was 10th (first year at high schools). Judging from age of acquisition for content words (WRDAOAc, no.19), Flesch reading ease (RDFRE, no.21) and Flesch-Kincaid Grade level (RDFKGL, no.22) the 2015 essay was easy and the 2016 essay was complex. As a result, it may be possible to say that the 2016 essay was coherent with a higher in readability than the 2015 essay. The gap between the 2015 essay and the 2016 essay was large in terms of grade level but the difference of TOEIC scores was only 20 points. With regard to TOEIC scores, the quality of writing was not affected by English proficiency assessed by TOEIC. The reason for the result of Coh-Metrix was that the 2016 essay had high cohesion with many connectives. As can be seen from these differences, with the author’s writing, the important thing might seem to be the knowledge about coherence and cohesion. The contents that the author had learned after the 2015 essay but prior to the 2016 essay were how to connect sentences, paying attention to coherence, and reviewing writing carefully. However, this analysis has limitations because both essays were written in a different situation, so the results might be difficult to be generalized.

2.6 Results and discussion of Theme-Rheme structures analysis for the 2015 essay and the 2016 essay

Table 2 Theme-Rheme structure of the 2015 essay

Theme		Rheme
Students	→	should choose whether study abroad or not
First all of students	→	forced to study ~ money.
The expense ~abroad	→	<u>too expensive</u>
<u>It's</u>	↙	
	→	unfair because ~ different.
some students	→	brought up in fatherless families

\*Excerpt from the 2015 essay.



Students should choose whether study abroad or not. First, all students are forced to study abroad and pay a lot of money. The expense of studying abroad is too expensive. It's unfair because parents' occupations and incomes are totally different. Moreover, some students are brought up in fatherless families.

Table 3 Theme-Rheme structure of the 2016 essay

Theme		Rheme
39 students	→	participated and took a speaking test
They	→	took as English Course at a SELHi
Cross-Sectional Study	→	<u>were approximately 40 students</u>
	↙	
<u>They</u>	→	took English~ type of instruction
Students	→	met an interviewer~ a picture

\*Excerpt from the 2016 essay.

In the Longitudinal Study, 39 students (20 male and 19 female) participated and took a speaking test three times. They had high motivations to learn and use English and took as English Course at a SELHi. In the contrast, Cross-Sectional Study participants were approximately 40 students in the first, second, and third years. They took English Course at the SELHi and received similar types of instructions. Students met an interviewer who asked the students to describe a picture for a few minutes.

The 2015 essay and the 2016 essay were examined by using Theme-Rheme structure. Each clause was divided into a Theme and a Rheme. Then, as for successive cross-referential links from Rhemes to Themes, a Rheme which became a Theme of the next clause was shown with an underline and an arrow in Table 2 and 3. Based on Theme-Rheme structures in Table 2 and Table 3, both texts lacked Thematic progressions. Each of the texts had just one cross-referential link. In the 2015 essay, the

Rheme of the third clause “too expensive” became the Theme of the next clause “it’s”. In the 2016 essay, the Rheme of the third sentence “were approximately 40 students” became the Theme “They” of the next clause. Both texts lacked Thematic progressions and ideas in the texts were not linked well. Therefore, from a Thematic perspective, both of the texts lacked coherence. It should be noted that the 2016 essay did not have enough cross-referential links even though it was found to have objectively high numbers of cohesive devices and high in readability according to Coh-Metrix.

### **3. Development of a Coherence and Cohesion Checklist and a List of Common Errors**

One of the purposes of this paper is to provide a solution to the issues of teaching coherence and cohesion. What instructional tools might help teachers? One solution to teachers’ unclear feedback pointed out by Lee (2002), might be a *checklist*. The checklist (Appendix 1) might aid communication between teachers and learners by providing a common language to talk about coherence and cohesion. It contains four items: Cohesion, Development, Theme and Rheme. Based on Halliday & Hasan (1976), Halliday (1985), and Eggins (1994), the checklist was created for classroom use in an EFL high school context. It is mainly for high school context because high school students have many opportunities to write long essays than junior high school students. As can be seen from the result of Coh-Metrix, cohesion improved in the 2016 essay, probably because the author learned how to connect sentences with connectives and adverbs such as “plus” and “thus”. Yet, coherence in terms of Theme-Rheme structure did not improve; both of the essays lacked Thematic progression. Therefore, acquisition of Theme-Rheme patterns may be more difficult, so learners might need clearer instruction. Common errors of discourse committed by inexperienced writers may help teachers detect problems of coherence more easily. For teachers and learners, *a list of common errors* (Appendix 2) gives explanations and examples in order to review their writing easily. In this paper, the four common errors: lack of Thematic progression, brand new Theme, double Rhemes, and empty Rheme are used. By

referring to Bloor and Bloor's (1992) categories and Eggins (1994), the list was created using examples from EFL writers. Examples of each problems were excerpted from Oi (2008, p.40), Belmonte & McCabe-Hidalgo (1994, p.24-p.25), Wang (2007, p.168) and the 2015 essay and the 2016 essay. A teacher may focus on correcting grammatical errors unless otherwise instructed. However, writing should be reviewed at the all level. Knowledge of Theme-Rheme structure and common errors would draw teachers' attention beyond surface errors to the discourse-related quality of writing.

The checklist and the list of common errors created by the author in this paper are intended to help teachers review learners' texts. These materials are especially for high school context. Compared to junior high school, high school students are required to write longer sentences and they have more opportunities to write essays. In particular, this checklist is targeted especially to new English teachers or learners who do not know how to review writing at the discourse level. Inexperienced writers have tendency to revise their writing only at the lexical level (Connor & Farmer, 1990). Teachers are also apt to pay attention to grammatical errors in learners' writing and may not know how to review learners' writing at the discourse levels. To improve quality of writing, writing needs to be revised at all levels. If teachers use the checklist and the list of common errors in class, learners would also be conscious of problems in their own writing. That is why the checklist and the list of common errors could draw not only teachers' but also learners' attention to the quality of writing. The checklist has four categories: 'Cohesion' , 'Theme' , 'Rheme' , and 'Development' . The description of each item is shown below.

### **3.1 Cohesion: Demonstratives, Pronouns, Connectives, Articles and Keywords**

#### **3.1.1 Demonstratives and Pronouns**

Demonstratives and pronouns are useful in avoiding writing the same words many times. Nouns and proper nouns can be replaced by other words such as "this", "that", "it", "she", "he" or "they".

#### **3.1.2 Connectives**

Connectives also known as conjunctions are classified into additive, adversative, causal and temporal. (Halliday & Hasan, 1976). Examples of each category are

“moreover”, “however”, “as a result” and “lastly” respectively.

### **3.1.3 Articles**

“A” and “an” are used with an indefinite noun such as a thing that is mentioned in a text for the first time. In contrast, “the” is used with a definite noun, which is already known. For example, “I watched *a* movie yesterday. *The* movie was not interesting.”

### **3.1.4 Enumeration**

This item is for checking whether there are enumeration including “first”, “second”, and “last” in texts.

### **3.1.5 Keywords**

Keywords including hypernyms, hyponyms and synonyms are regarded as important elements of cohesion. A synonym is a word with a similar meaning to another word. For example, the nouns, “help” and “aid” have similar meanings. They are said to form what is known as lexical chain. The lexical chain can enhance cohesion which will then contribute to topical development. A lexical chain containing an abstract hypernym and concrete hyponyms would not only create surface cohesion but enhance coherence at a deeper level by developing the topic. With these devices, texts would be more coherent.

## **3.2 Development**

This item is for examining whether the text is written in accordance with the topic, and whether a Rheme offers new information to further develop the Theme. The topic of the text should be reflected in the Theme positions. Rhemes should offer new information to develop Themes. Some learners’ writing lack Thematic progression, resulting in “I poems”. The writer may write about the same Theme with completely unrelated Rhemes in every sentence.

## **3.3 Theme and Rheme**

The items in the checklist for Theme and Rheme mainly focused on brand new Theme and empty Rheme. As mentioned earlier, Theme has old or familiar information. Rheme has new or unfamiliar information. Additionally, a Theme tends to be based on the Rheme of the previous clause. In other words, a Rheme tends to

become the Theme of the next clause. To see the relation between Themes and Rhemes, Theme-Rheme structure (Table 2 and 3) would be informative.

#### **4. Conclusion**

This paper revealed problems in one EFL learner's coherence and cohesion, and provided a potential solution in the form of instructional material. In order to see changes in the author's writing, this study used Coh-Metrix and Theme-Rheme structure. The results revealed that the author's use of cohesive devices (some connectives and parts of speech incidences) improved only slightly despite the fact that a large difference was found in the readability levels. The author's English proficiency between the two essays was not improved well, but the objective quality of writing was improved. In terms of Theme-Rheme structure, however, both of the essays lacked Thematic progressions. Even the 2016 essay had only one cross-referential link. From the improvements in the author's essays, knowing how to connect ideas might be an important factor in improving coherence and cohesion. It would be advisable that the teacher focuses on teaching how to connect ideas by applying principles of coherence and cohesion. That is, the information structure of the sentence needs more focus, perhaps even more so than the hamburger-style organization. Connecting ideas and sentences is an important aspect of writing as well as organizing ideas. And also, the checklist (Appendix 1) and the list of common errors (Appendix 2) would be helpful to review learners' writing at all levels. The course of study (MEXT, 2009) does not provide concrete examples to show teachers how to teach coherence and cohesion in the classroom. That is why, as a learner, the author had never had the opportunity to learn about coherence and cohesion especially the way of connecting ideas until meeting an American instructor in university. Lastly, this paper has limitation, and it remains a need for further research. The results of Coh-Metrix and Theme-Rheme structure analysis have difficulties to be generalized because the 2015 essay and the 2016 essay were written in different situations. Further research is necessary to analyze essays written in that have the same setting or to the same

prompts. The checklist and the list of common errors might become one solution to the problem of teachers' vague feedback on coherence in learners' writing which Lee (2002) pointed out. The author feels that the checklist would help teachers detect problems in learners' writing. Furthermore, it might enable teachers and learners to have better communication about coherence and cohesion with ease. However, effectiveness of the checklist and the list are not tested in classroom yet. When the author starts working as an English teacher next year, the practicality of the checklist and the list of common errors would be tested and revised in future research.

## References

- Bain, A. (1890). *English composition and rhetoric*. Longmans, Green & Company.
- Belmonte, I. A., & McCabe-Hidalgo, A. (1998). Theme-Rheme patterns in L 2 writing. *Didáctica*, 10, 13-31.
- Benesse Educational Research and Development Institute. (2009). *The first research of the report study on lower secondary school English*. Retrieved from <http://berd.benesse.jp/global/research/detail1.php?id=3186>. (Accessed 2018-1-31)
- Bloor, M., and Bloor, T. (1992). Given and new information in the thematic organization of text: An application to the teaching of academic writing. *Occasional Papers in Systemic Linguistics*, 6, 33-43.
- Coh-Metrix (version 3.X). Available from <http://cohmetrix.com>. (Accessed 2018-1-27)
- Connor, U., & Farmer, F. (1990). The teaching of topical structure analysis as a revision strategy for ESL writers. In B. Kroll (Ed.), *Second Language Writing: Research Insights for the Classroom*. (pp. 126-139) Cambridge University Press.
- Eggins, S. (1994). *An introduction to systemic functional linguistic*. London: Printer Publishers.
- Fries, P. H. (1983). On the status of theme in English: Arguments in discourse. In J.S. Petofi & E. Sozer (Eds.), *Micro and macro connexity of text*. Hamburg: Helmut Buske Verlag.

- Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. London: Longman.
- Halliday, M. A. K. (1985). *An Introduction to Functional Grammar*. London: Edward Arnold.
- Kudo, Y. (2011). The Feature of Cohesion in Japanese English Learners' Writing. (Doctoral dissertation, Tokyo University of Foreign Studies).
- Lee, I. (2002). Teaching coherence to ESL students: A classroom inquiry. *Journal of second language writing*, 11 (2), 135-159.
- Ministry of Education, Culture, Sports, Science and Technology. (2008). *The course of study for foreign language for lower secondary school*. Retrieved from [http://www.mext.go.jp/component/a\\_menu/education/micro\\_detail/\\_\\_icsFiles/afldfile/2017/07/25/1387018\\_10\\_1.pdf](http://www.mext.go.jp/component/a_menu/education/micro_detail/__icsFiles/afldfile/2017/07/25/1387018_10_1.pdf). (Accessed 2018-2-5)
- Ministry of Education, Culture, Sports, Science and Technology. (2009). *The course of study for foreign language for Upper secondary school*. Retrieved from [http://www.mext.go.jp/component/a\\_menu/education/micro\\_detail/\\_\\_icsFiles/afldfile/2010/01/29/1282000\\_9.pdf](http://www.mext.go.jp/component/a_menu/education/micro_detail/__icsFiles/afldfile/2010/01/29/1282000_9.pdf). (Accessed 2018-2-5)
- National Institute for Educational Policy Research. (2005). *The course of study implementation survey of elementary and lower secondary schools*. Retrieved from [https://www.nier.go.jp/kaihatsu/katei\\_h15/H15/03001051030007004.pdf](https://www.nier.go.jp/kaihatsu/katei_h15/H15/03001051030007004.pdf). (Accessed 2018-2-5)
- National Institute for Educational Policy Research. (2007). *The course of study implementation survey of upper secondary schools*. Retrieved from [http://www.nier.go.jp/kaihatsu/katei\\_h17\\_h/h17\\_h/result\\_q511.pdf](http://www.nier.go.jp/kaihatsu/katei_h17_h/h17_h/result_q511.pdf). (Accessed 2018-2-5)
- Oi, K. (2008). *An Introduction to Paragraph Writing Guide: For Effective Writing Guide at Junior and Senior High School*. Tokyo: Taishukan.
- Wang, L. (2007). Theme and Rheme in the thematic organization of text: Implications for teaching academic writing. *Asian EFL Journal*, 9 (1), 164-176.

Appendix 1. Checklist

Categories	Lists	Check
Cohesion	<u>Demonstratives (this, that, it) and pronouns (he, she, they) are effectively used.</u> └ 指示語や代名詞は使われている。	
	<u>Connectives (because, after, if, although) are effectively used.</u> └ 接続詞（理由、時制、仮定、逆説）は使われている。	
	<u>Articles (a, an, the) are used in the text properly.</u> └ a, an と the を正しく使い分けている。	
	<u>Enumeration (first, second, last) is done properly in the text.</u> └ 文章の中で数的列挙はされている。	
	<u>Keywords (hypernyms, hyponyms, synonyms) are repeated in the text.</u> └ 文章の中で重要語（上位語、下位語、同義語）が繰り返されている。	
Development	<u>The topic of the essay is reflected in the Themes.</u> └ 主題がテーマに反映されている（ずっと I で始まっている）。	
	<u>Rhemes develop Themes.</u> └ レーマがテーマを発展させている。	
Theme	<u>Themes have old or familiar information.</u> └ 節の中のテーマは古い、既知の情報である。	
	<u>Themes refer to Rhemes of the previous clause.</u> └ 前の節のレーマが次の節の中のテーマになっている。	
Rheme	<u>Rhemes have new or unfamiliar information.</u> └ 節の中のレーマは新しい未知の情報である。	



Appendix 2. List of common errors

Categories	Explanations and Examples
Lack of Thematic progression	<ul style="list-style-type: none"> <li>– The Rhemes do not develop the Themes (I-poem) .</li> <li>– レーマがテーマを発展させず、主題に沿って文章が書かれていない。</li> </ul> <p>e.g., <u>“I like soccer. I like Queen. I like Matsui. I like Summer.”</u> (Oi, 2008, p.40)</p>
Brand New Theme	<ul style="list-style-type: none"> <li>– New or unfamiliar information ('special unit', and 'a higher education') is mentioned in the Theme position.</li> <li>– テーマが新しく、未知の情報である。</li> </ul> <p>e.g., <u>“First special unit wasted our time.”</u> (excerpt from the 2015 essay)</p> <p>e.g., <u>“A higher education is a right (...).”</u> (Belmonte &amp; McCabe-Hidalgo, 1998, p.25)</p>
Double Rhemes	<ul style="list-style-type: none"> <li>– There are two Rhemes and the development of the Theme is not appropriate.</li> <li>– レーマが1つの節に2つあり、その内の1つが今まで述べられていないものである。</li> </ul> <p>e.g., <u>“The educational reform had a big influence on young teachers and the students’ families paid a lot of money for their children.”</u> (Wang, 2007, p.168)</p> <p>e.g., <u>“They [the students participated in the study] took English Course at SELHi and received similar types of instruction.</u> (excerpt from the 2016 essay)</p>
Empty Rheme	<ul style="list-style-type: none"> <li>– The Rhemes ('were vain effort', and 'a huge question') fail to offer new information.</li> <li>– レーマが新しく、未知の情報ではない。</li> </ul> <p>e.g., <u>“To make posters were vain effort.”</u> (excerpt from the 2015 essay)</p> <p>e.g., <u>“Nowadays, higher education is a huge question (...).”</u> (Belmonte &amp; McCabe-Hidalgo, 1998, p.24)</p>

### Appendix 3. Detailed description of labels of Coh-Metrix

No.1-paragraph count

No.2-sentence count

No.3-word count

No.4-the average number of sentences in a paragraph

No.5-lexical diversity type-token ratio in all words

(Type-token ratio is the number of unique words. For example, if the word cat appears 7 times in a text, the type value is 1. If the sentence consists of all different words, the lexical diversity is maximum.)

No.6~No.12-incidences of all connectives

No.13-the repetition score for tense and aspect

No.14~No.18- incidence (happening per 100 words)

No.19-age of acquisition (ex. milk, smile, and pony=202 dogma and matrix=700)

No.20-the score shows familiarity of a word for an adult (ex. milk=588, mother=632, calix=124)

No.21-the score shows easability (average text is 6 to 70, higher scores show easier reading)

No.22-the score shows grade-level (range is from 0 to 12)